DOCUMENT RESUME

ED 117 818 EA 007 874

AUTHOR Auster, Ethel; Lawton, Stephen B.
TITLE Automated Retrieval of ERIC Materials: Implications for Canadian Educators.
PUB DATE Jun 74
NOTE 16p.; Paper presented at the Annual Meeting of the Canadian Society for the Study of Education (Toronto, Ontario, June 6-8, 1974)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Bilingualism; *Data Bases; Educational Research; *Information Processing; *Information Retrieval; *Information Systems
IDENTIFIERS *Canada

ABSTRACT

The disadvantages for Canadians using the ERIC system are (1) fledgling Canadian computer information systems are deprived of their financial support; (2) Canadian research included in ERIC can not necessarily be retrieved by country of origin; (3) of the 200 journals indexed in Canadian Education Index (CEI), only 26 are to found among the 600 in ERIC's Current Index for Journals in Education (CIJE); and (4) policies regarding the inclusion, exclusion, and manner of indexing documents are beyond Canadian control. Solutions to some of these problems include (1) more systematic and exclusive utilization of the ERIC system, (2) creation of a Canadian "clearinghouse" indexing all Canadian documents for ERIC, and (3) creation of a bilingual Canadian data collection and indexing system. Considering the complexity of the situation involved in creating an all-Canadian system, interim dependence is deemed necessary for one to three years on the ERIC data base searched via an existing U.S. online retrieval system. (MLF)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

ED 078178

AUTOMATED RETRIEVAL OF ERIC MATERIALS: IMPLICATIONS FOR CANADIAN EDUCATORS*

Ethel Auster and Stephen B. Lawton
The Ontario Institute for Studies in Education

Gaining quick access to needed information is a perennial problem encountered by most researchers, teachers and administrators involved in the educational enterprise. Until recently, when a person lacked information, he would generally turn directly to libraries, journals, or individuals who had first-hand knowledge of the subject (Fry, 1972) in order to satisfy his needs. However, in the past year or two, still another mode of access has become practicable: automated on-line bibliographic searching. Judging from the rapid increase in the number of these services (ERIC Clearinghouse on Educational Management Newsletter, Winter 1974) and the extent of reported use of some of them (Spuck, 1974; Auster and Lawton, 1973), it appears that they are quickly becoming a major channel of access to educational literature. Because all of these services within the field of education depend upon the data base prepared by the Educational Resources Information Center (ERIC) in the United States, and because much of the computer hard-ware and soft-ward utilized to search the data is located in the U.S., Canadian educators are faced with the temptation of using efficient computerized methods of acquiring information from a source which may not include Canadian materials vital to their needs and, by their use of U.S. facilities, may deprive fledgling Canadian computer information systems of their financial support. The development of possible solutions to

*Paper presented at the Annual Conference of the Canadian Society for the Study of Education, Toronto, Ontario; June 6-8, 1974.

these problems is the topic of this paper.

There are still those, of course, who question whether automated systems of gaining access to printed information are necessary. But, as Anderla (1973) has noted, "Information threatens to become a source of pollution rather than greater wealth...[We must] learn to channel the generation of this information wisely [and]...to organize its distribution....Massive automation will come about spontaneously under the irresistible thrust of information's exponential growth and technological progress (pp. 27-28)."

In essence, the task of meeting the demands of those who need or desire current information falls into three phases: the generation of information, its collection and indexing, and its retrieval for a given purpose. The second and third stages--assembly of the data base and retrieval of specific items represent the challenge that must now be faced. Unfortunately, there is no adequate Canadian data base currently available to which one can gain access by using an automated retrieval system. Nor does there appear to be a fully developed Canadian retrieval system that could be used to search such a data base.

To be sure, there are indexes to much of the Canadian educational literature, notably the Canadian Education Index(CEI) published by the Canadian Education Association. However, it is limited in that it does not index the vast quantity of "fugitive" report literature, such as project reports written for administrative purposes within school boards and which are not published, although CEA does publish an annual Directory of Education Studies in Canada which includes a part of this information. In addition, CEA has recently initiated steps toward

capturing some of this fugitive literature in microform to facilitate its dissemination. However, CEI and the Directory are not currently produced in machine-readable fashion allowing for automated retrieval. For a data base containing Canadian material that can be searched by machine, one must turn to ERIC.

Detailed descriptions of the ERIC system are readily available (e.g., Mathies and Watson, 1973; Fry, 1972; Spuck, 1974), but a brief overview will be included here for those unfamiliar with the system. The chief objectives of ERIC, according to Fry (pp. 6-7), are

1) to make significant but previously unavailable R & D reports easily and readily available,

2) to interpret and summarize results in ways that educational practitioners and decision makers can use them,

3) to help strengthen existing educational communication channels for putting R & D results to use, and

4) to become and important base for developing a national information system [for the United States].

The major structural features of the organization to carry out these tasks evolved over a six year period from 1959, when the United States Office of Education, faced with mounting numbers of reports "literally piling up in the halls (Spuck, p. 2)" which "thereafter became hard to find and later even difficult to identify (ibid)," supported a feasibility study for an in-house information system. The broader utility of such a system to other agencies concerned with education became apparent and in 1965 funds were appropriated as part of the Elementary and Secondary Education Act (ESEA) for what eventually became ERIC.

A decentralized structure was adopted for ERIC, including a number of clearinghouses specializing in different subject areas and coordinated by a central office now located at the National Institute of Education (NIE). While each of the sixteen clearinghouses now in existence specialize in such areas as career education, reading, teacher education, and educational management, and while each has considerable autonomy, terminology is controlled by use of the authoritative Thesaurus of ERIC Descriptors (CCM, 1972.)

Insert Figure 1 about here

In practice, the various clearinghouses have the responsibility of acquiring, screening, indexing and abstracting fugitive documents. The final collection is housed centrally, and reports can be obtained either in microfiche or hard-cover from the ERIC Document Reproduction Service (EDRS) operated by LEASCO corporation. The index for reports is published as Research in Education (RIE), which is used for manual searches involving authors or subject descriptors, the latter being selected from the ERIC Thesaurus. Since 1969, journal articles have also been indexed by the clearinghouses; these are reported in the Current Index for Journals in Education (CIJE).

In addition to being published in RIE and CIJE, the bibliographic information for ERIC is available in the form of magnetic, machine readable data tapes referred to as ERICTAPES. These can be searched using appropriate combinations of computers (hard-ware) and programs (soft-ware).

The size of the ERIC data base, including bibliographic references to reports in RIE and to journal articles in CIJE, now exceeds 160,000 items

as is shown graphically in Figure 2. Needless to say, part of the attraction of automated searching is the reduction in time and increase in effectiveness which can be achieved over manual searching. At the same time, the number of items is growing constantly, with new volumes of RIE and CIJE being published monthly. Fortunately, ERICTAPES are updated with equal frequency.

Insert Figure 2 about here

In addition to the growth in the data base itself, there has been an increase in the number of descriptors used to classify documents.

Insert Figure 3 about here

To be sure, the rate of increase in the numbers of descriptors is leveling off to perhaps 100 per year (Spuck, 1974), but continued additions are inevitable because education is an interdisciplinary subject which uses a broad range of terms, most of which have many facets and few of which have precise definitions. Also new terms will be invented and old ones given new meanings.

Although ERIC is a national information service for education in the United States, it does not restrict itself to the collection of U.S. materials and, in fact, contains significatn amounts of Canadian materials. The extent of the Canadian content of ERIC has been investigated in several recent studies (Auster and Lawton, 1973 and Summers, 1974) but precise conclusions are exceedingly difficult to make since references cannot be retrieved on the basis of national origin or on the geographical location in which a given study took place. Never-

theless, use of the ERIC Institution Index (Macmillan Information, 1974) in executing searches revealed, for example, sixty references carrying the code for the Ontario Institute for Studies in Education (OISE)—which obviously represents only a small portion of the work originating there.

Because the ERIC data base in the form of ERICTAPES can be searched conveniently using an on-line computer system such as SDC/ERIC offered by the System Development Corporation, a batch search system like that employed by the National Science Library for its Selective Dissemination of Information (SDI) program, or even by doing manual searches in university and board libraries, it can probably be considered the data base most readily available to Canadian educators, particularly since complete ERIC microfiche collections are maintained in a score of locations across Canada (Fry, 1972). But, though ERIC may be useful, is it entirely adequate for Canadian needs? If not, what are the characteristics that an ideal Canadian system should have? In answering these questions, consider the latter one first.

It is our belief that an effective information system for Canadian education must have at least these five qualities:

1. Comprehensiveness. The collection must be as nearly inclusive of all printed materials produced in Canada of sufficient quality and content to be of use and interest to practising educators and researchers. If there are gaps in a collection, then one may unknowingly overlook references that would be of great value. Further, after repeated disappointments created by missed references, a user would probably lose confidence in the system and cease to use it.

2. Develop Canadian expertise. Creation and maintenance of the data file would create jobs and develop expertise within Canada. These

would serve to guarantee the relevance and accuracy of indexing, and to ensure control of policies affecting the collection's quality, content, and future.

3. Adequate indexing. An adequate system of indexing must be used, lest important works become irretrievable by being indexed under illogical, irrelevant, incorrect or hopelessly vague terms. Precision is needed. In addition, a bilingual French and English system is to be preferred.

4. Currency. The collection must be up-to-date. Materials must become available in the collection within a few months of their completion. More than a quarterly delay should be avoided, both because issues in education arise quickly and because researchers must be able to ensure that their work does not duplicate that of others.

5. Modest cost. Both the cost per unit and total cost of assembly, maintenance and reproduction of the data base must be sufficiently low to be born by Canadian educational institutions and appropriate financing agencies.

ERIC fails to meet at least the first two criteria--comprehensiveness and national development--entirely, and does not completely fulfill the third--adequate indexing. Only in its currency and cost is it totally acceptable. That ERIC does not contain extensive amounts of Canadian materials is perhaps its most serious shortcoming. In contrast to CEI which indexes about 200 Canadian journals, only 26 are to be found among CIJE's 600. Also, relatively few reports find their way from Canada to the ERIC clearinghouses for abstracting in RIE. In general, the process of submission and acquisition of Canadian documents appears haphazard, depending on the habits of a relatively few individuals who have chosen to send their work to the clearinghouses. The fact ERIC

Central will now accept documents, redirecting them to the appropriate clearinghouse, makes this task easier and may serve to increase the rate of submission, but it is too soon to judge the effects of this recent policy. In any case, ERIC has a weakness more fundamental than its lack of Canadian documents: policies regarding the inclusion, exclusion and manner of indexing documents are beyond Canadian control. A document relevant to Canadian education might easily be refused if it were irrelevant to American education. Terminology, which differs between the U.S. and Canada, does not include many terms used only in Canada, or used with different meanings. Yet, because of its convenience and wide availability, dependence on ERIC appears to be growing in Canada, perhaps to the detriment of Canadian institutions, publications, and expertise.

What might be done to solve these problems? There appear to be three alternatives that warrant serious consideration at this time--both in terms of the immediate situation and of long term objectives.

The first possibility is more systematic and extensive utilization of the ERIC system. While not overcoming the objections regarding the locus of policy decisions and the failure to enhance Canadian expertise, this alternative would ensure more extensive inclusion of Canadian material. It might also be possible to persuade ERIC to adopt more Canadian terms, and to use "Canada" as a description for all items originating in or pertaining to Canada. But, because of the negative effects of depending almost solely on ERIC, particularly over a long period of time, this approach can probably be considered at most a practical interim solution until another choice is readily available.

Creation of a Canadian "clearinghouse" which, rather than being subject oriented, would index all Canadian documents for ERIC would be

a second alternative. Such a clearinghouse would not conform to the ERIC model, however, and might better function as a device for initial screening, indexing, and abstracting of documents later forwarded to the appropriate ERIC subject specialty clearinghouse for review and final processing. Such a proposal would require the cooperation and agreement of ERIC, perhaps with guarantees for acceptance by ERIC of all documents recommended by the Canadian "preclearinghouse", the regular clearinghouse having authority only to add descriptions and appropriate identification codes. A quota could be set in such an agreement to protect ERIC from a flood of Canadian documents. This alternative would provide greater support for the development of expertise in indexing, abstracting, and evaluating documents within Canada. Also, a major part of the decision making powers would be concentrated here; Canada would carry an equitable share of the cost of the ERIC system, which would not be true if totally unprocessed documents were sent to ERIC; and a unified educational information system would be created for North America. Yet, this alternative would be open to valid criticisms as being in the tradition of "branch plant" operations wherein U.S. companies export their technology to Canadian subsidiaries while retaining all research and development projects at home. If implemented permanently, such a Canadian preclearinghouse might retard development of bilingual indexing techniques now being developed.

The final, and perhaps most attractive alternative, would be creation of a bilingual Canadian data collection and index. Creating a bilingual index which would conform completely to ERIC is probably impossible, so this data base would be less comprehensive than an augmented ERIC system (though ERIC would still be available for separate

searches). On the other hand, francophones would not have to search in English, as is true in ERIC which includes some references to French material under English translations of titles. The greater complexity of bilingual indexing would probably make monthly up-dates unrealistic, though quarterly up-dates could probably be achieved. The cost of such a system would exceed that of the other alternatives; yet, the approach would support the developing expertise in bilingual indexing currently found in Quebec at such places as Universite de Laval and the Centre for Animation, Development, and Research in Education (CADRE).

Retrieving references and abstracts from the data base is the other activity that must be considered, complementing as it does creation and maintenance of the data base. Since comparison of on-line and batch systems invariably favour the former (Auster and Lawton, 1973), it is reasonable to assume that on-line systems will prevail, especially in view of recent reductions in the cost of time-sharing on computers.

The work in on-line retrieval is done via a computer program—the software. Software can be "rented", in the sense that searches are done on a per-use basis on someone else's system (e.g., SDC/ERIC); software can also be purchased outright and put up on one's own computer (e.g., SDC will sell its ORBIT II search program for $40,000-$50,000); or a program can be written "in-house" (e.g., the National Science Library's experimental CAN/OLE program and Queen's QUIK-LAW or QL).

The relevant criteria to be applied to selection of soft-ware, or the use thereof, include (1) cost-effectiveness, (2) availability, and (3) effect on Canadian development. Obviously, the choice one makes depends on the organizational context. A relatively small organization

wishing to initiate searches of the ERIC data base immediately would, on the basis of cost-effectiveness and availability, select the first alternative and open an account with a search service. If, however, a large province were to commit itself to using an extensive network of computer terminals to query the data base, as the State of Wisconsin has done (Spuck, et al, 1974), then the second choice is feasible. If, and only if, a total Canadian commitment were made could the investment--both in funds and time--to create a comparable Canadian software package probably be justified. In any case, interim use of a U.S. system would still be an attractive option.

Given the various alternatives concerning data bases and search systems, and considering the complexity of the situation with a large number of parties being involved in decisions to create an all-Canadian system we are inclined to accept interim dependence, for 1 to 3 years, on the ERIC data base searched via an existing U.S. on-line retrieval systems. This resource could be supplemented by a drive to increase the number of Canadian contributions and descriptors in ERIC. But whether this interim period should be used to create one or several Canadian clearinghouses for ERIC, or to create a totally Canadian system, is a decision that must be made at a national level--by the Council of Ministers of Education, the National Library, the Social Science Research Council, the Canadian Society for the Study of Education, or a coalition of these and/or other bodies. The decision to act must not be delayed. It is our hope that this session will provide the impetus for action in the near future.

REFERENCES

Anderla, G. "The Future of Information: A Challenge for Governments and Society." OZCD Observer, April, 1973; pp.27-32.

Auster, E. and Lawton, S.B. Meeting Ontario's Need for Educational Information: An Evaluation of the SDC/ERIC On-line Bibliographic Search Service. Toronto: The Ontario Institute for Studies in Education. (An informal publication of the Library, OISE), 1973. Available as ED 077 531.

Canadian Education Association, Canadian Education Index. Toronto: CEA (quarterly).

Canadian Education Association, Directory of Education Studies in Canada. Toronto: CEA (annual).

Educational Resources Information Center, Current Index to Journals in Education. New York: Macmillan Information, (monthly).

Educational Resources Information Center, ERIC: A Profile. Washington, D.C.: ERIC, National Institute for Education, 1974.

Educational Resources Information Center, Research in Education. Washington, D.C.: U.S. Government Printing Office, (monthly).

Education Resources Information Center, Thesaurus of ERIC Descriptors. New York: CCM (now Macmillan Information), 1972.

ERIC Clearinghouse on Educational Management, Newsletter. Eugene, Oregon: ERIC Clearinghouse on Educational Management, March, 1972.

ERIC Institution Index. New York: Macmillan Information, 1974.

Fry, B.M. Evaluation Study of ERIC Products and Services, Summary Volume. Bloomington: Graduate School Library, University of Indiana, 1972. Avaliable as ED 060 922.

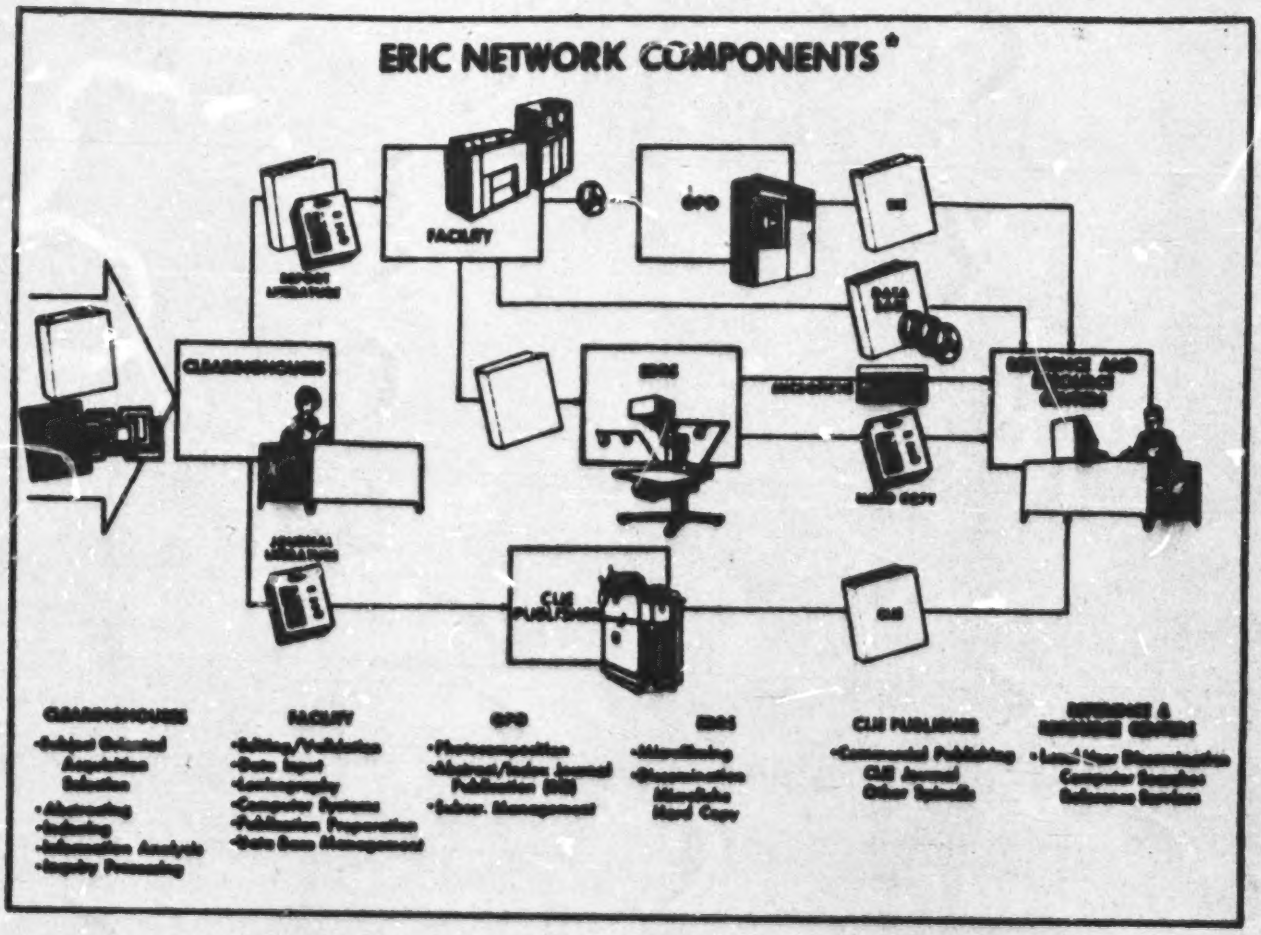
Mathies, M.L. and Watson, P.G. Computer-Based Reference Service. Chicago: American Library Association, 1973.

Spuck, D. "The ERIC System: History and Analysis." Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois; April 19, 1974.

Spuck, D.W., McIsaac, D.N., Olson, T. and Tally, R. "Information Retrieval: Presentation and Demonstration of an Interactive Computer-based Search Program." A symposium presented at the Annual Meeting of the American Educational Research Association; Chicago, Illinois; April 19, 1974.

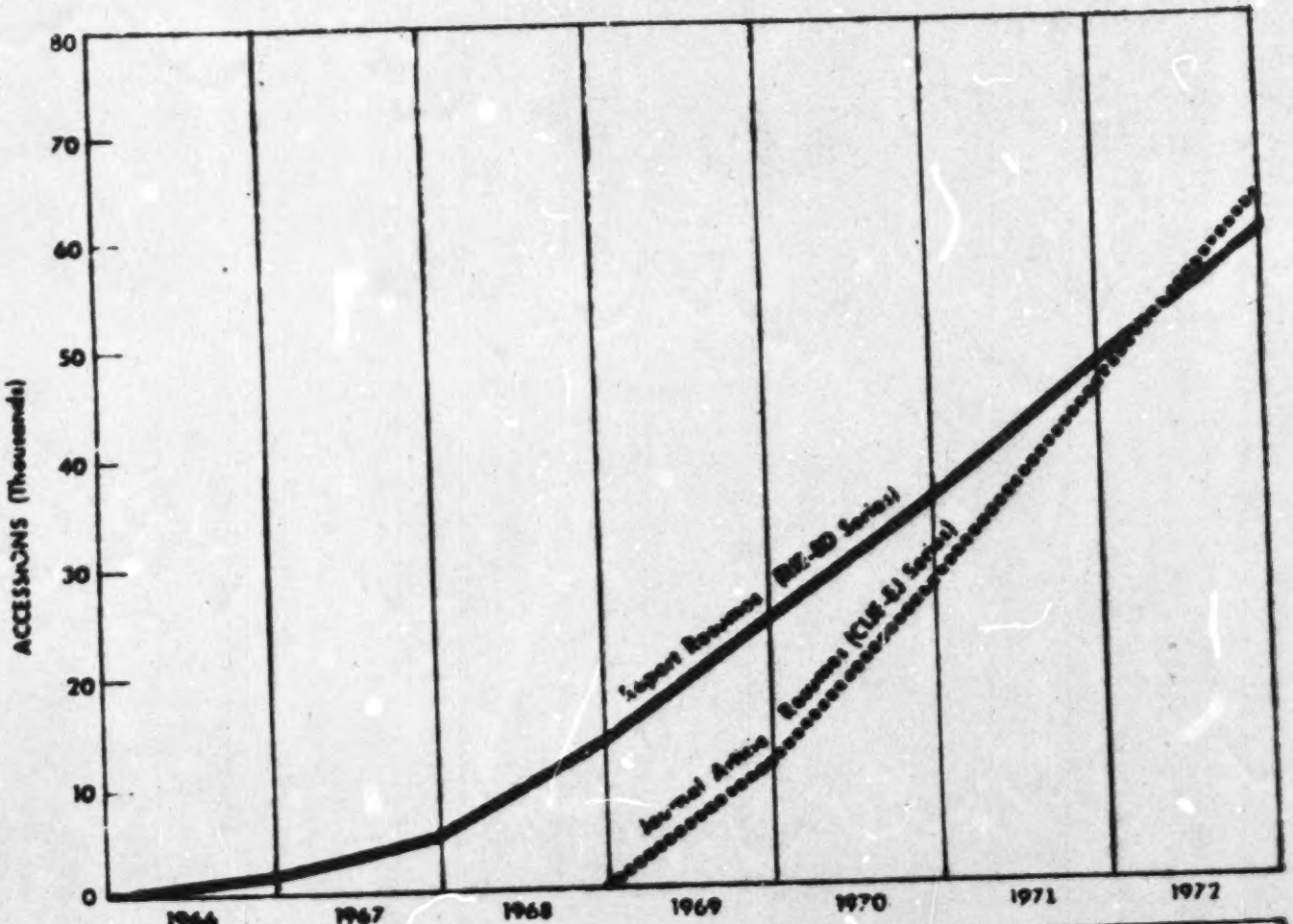
Summers, E.G. "Information and Bibliographic Needs in Canadian Education." A paper prepared for the National Conference on the State of Canadian Bibliography, University of British Columbia, May 22 - 24, 1974.

FIGURE 1



*ERIC: A Profile. Washington: NIE, 1974.

ERIC DATA BASE-FILE GROWTH

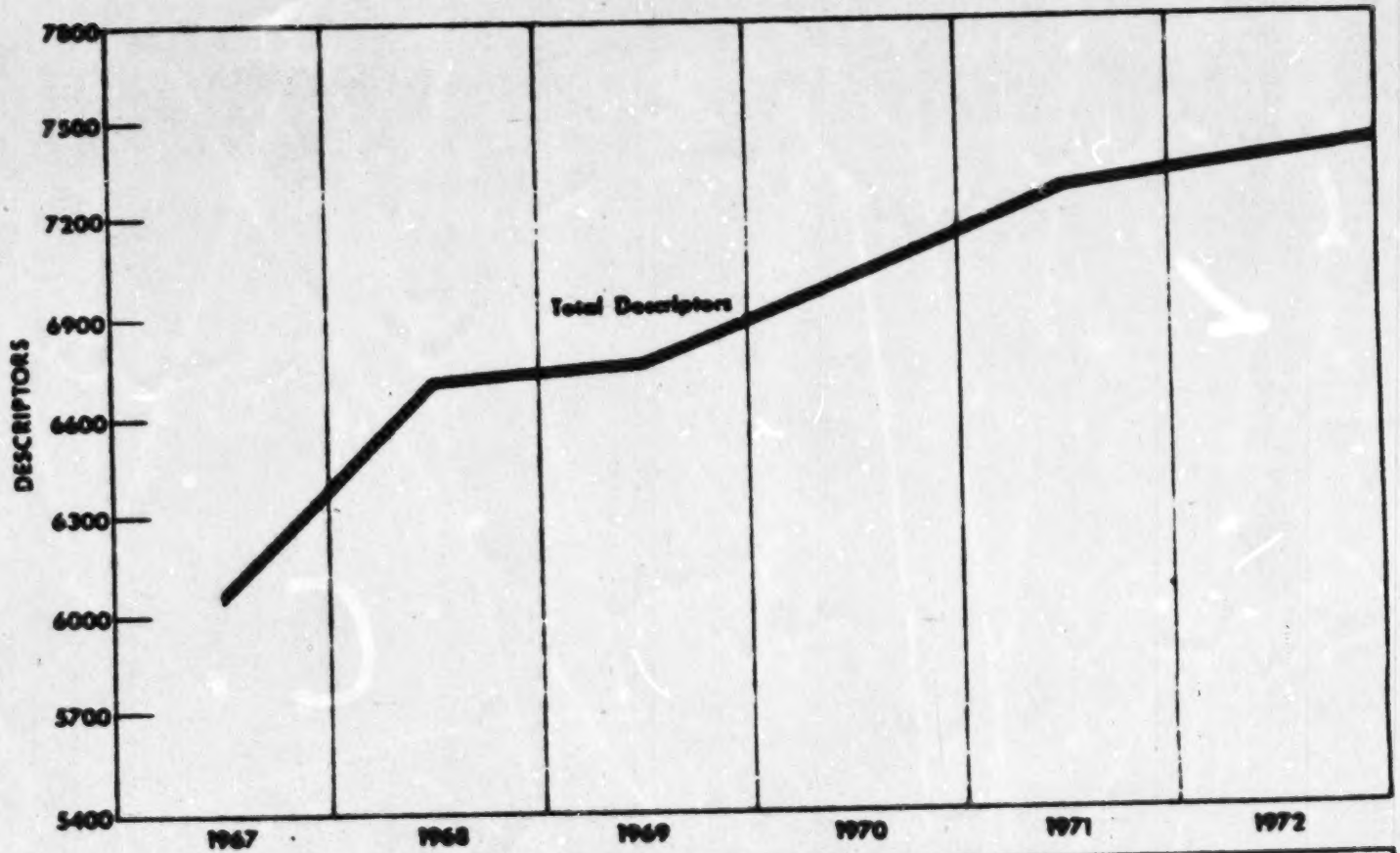


Reports	Accessions Added	1834	3469	8803	10,453	10,456	12 330	12 230
	Cumulative Total		5303	14,106	24,559	35 015	47 345	59 575
Articles	Accessions Added				11,707	15,892	17 672	17,480
	Cumulative Total				11,707	27,599	45,271	62,751

* ERIC Data Base Users Interchange.
February, 1973.

FIGURE 3

ERIC THESAURUS-FILE GROWTH*



	1967	1968	1969	1970	1971	1972
Use References	1126	1847	2140	2208	2395	2443
Main Terms	3411	4239	4562	4761	4877	4935
Total Descriptors	4537	6086	6702	7069	7272	7378

* ERIC Data Base Users Interchange.
February, 1973.